Between-day reliability of IMU-derived spine control metrics in patients with low back pain

Graham, Ryan B.; Dupeyron, Arnaud; van Dieën, Jaap H.

published in
Journal of Biomechanics
2020

DOI (link to publisher)
10.1016/j.jbiomech.2020.110080

document version
Publisher's PDF, also known as Version of record

citation for published version (APA)
Graham, R. B., Dupeyron, A., & van Dieën, J. H. (2020). Between-day reliability of IMU-derived spine control metrics in patients with low back pain. *Journal of Biomechanics*, *113*, 1-7. [110080]. https://doi.org/10.1016/j.jbiomech.2020.110080

Journal of Biomechanics 113 (2020) 110080

Contents lists available at ScienceDirect

Journal of Biomechanics

journal homepage: www.elsevier.com/locate/jbiomech
www.JBiomech.com

Between-day reliability of IMU-derived spine control metrics in patients with low back pain

Ryan B. Graham [a,*], Arnaud Dupeyron [b,c], Jaap H. van Dieën [d]

[a] School of Human Kinetics, Faculty of Health Sciences, University of Ottawa, 200 Lees Avenue, Ottawa, Canada
[b] EuroMov Digital Health in Motion, Univ Montpellier, IMT Mines Ales, 34090 Montpellier, France
[c] Department of Physical and Rehabilitation Medicine, CHU Nimes, Montpellier University, Nimes, France
[d] Department of Human Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam Movement Sciences, 1081 BT Amsterdam, the Netherlands

ARTICLE INFO

Article history:
Accepted 12 October 2020

Keywords:
Wearable sensors
Local divergence exponents
Variability
Spine
Neuromuscular control

ABSTRACT

Inertial measurement units (IMUs) are a potentially useful tool for clinicians and researchers in assessing spine movement biomechanics and neuromuscular control patterns. This study assessed the between-day reliability of the HIKOB FOX IMU in measuring local dynamic stability (LDS) and variability of trunk movements in patients with chronic low back pain (LBP). The local divergence exponent (λ_{max}) was used to quantify LDS and the mean standard deviation (MeanSD) between cycles was used to quantify variability during 30 repetitive cycles of flexion/extension, rotation, and complex movement tasks. For λ_{max} the average coefficient of variation (CV) was ~10% in the flexion/extension and rotation tasks, and all CV values were <20% when also including the complex task. ICC values for λ_{max} ranged from 0.28 to 0.81. Reliability of λ_{max} was similar between the pelvis and thorax segments (CV: ~10%, ICC: 0.48–0.78) and worse for the lumbar spine (CV: ~15%, ICC: 0.28–0.59). The CV for MeanSD was typically in the range of 20–30%, with even greater CV in the non-primary axes during each task (30–52%). Similarly, ICC values were lowest about the anterior-posterior axis in the flexion/extension task (ICC: 0.15–0.29) and largest about the longitudinal axis in the rotation task (ICC: 0.76–0.88). The moderate between-day reliability of λ_{max} in the sagittal and transverse planes offers improvement over manual and subjective tests with poor reliability that are currently used in clinics. The minimal detectable differences presented give a threshold for change in research and rehabilitation in patients with LBP.

1. Introduction

Low back pain (LBP) affects over 80% of people at some point in their lifetime (Andersson, 1999). Of these cases, 85–90% are classified as non-specific (Waddell, 2005), meaning that the pain is not attributed to any specific injury or pathology (Balagué et al., 2012). LBP places a large socioeconomic burden on health care systems worldwide. For instance, when considering premature mortality, prevalence and condition severity, the burden of LBP is second only to heart disease in the European Union (Murray et al., 2012), and is still the first cause of years lived with disability worldwide (Wu et al., 2020).

Despite the high prevalence and impact of LBP, overall assessment and management of the disorder is substandard, which does not permit clinicians to categorize patients or prescribe optimal individual treatment. LBP disorders have been documented to manifest in terms of altered movement quality, including excessive or poor stability and poor coordination of lumbopelvic segments (Asgari et al., 2015, 2017; Mokhtarinia et al., 2016; Seay et al., 2011; Spinelli et al., 2015). Local dynamic stability (LDS) estimates – whereby the average exponential rate of divergence between nearest neighbor trajectories in a reconstructed state space is used to estimate the amount of chaos in a system (Rosenstein et al., 1993) – have also been used to evaluate spine control and movement stability (Asgari et al., 2017; Bourdon et al., 2019; Graham et al., 2014; Granata and England, 2006; Granata and Gottipati, 2008; Ross et al., 2017). The local divergence exponent (λ_{max}), also known maximum finite-time Lyapunov exponent, is used to quantify LDS and is a metric that could be used to establish subgroups of LBP patients, with those with high λ_{max} values having "loose" control of the spine and possible proprioceptive deficits and those with low λ_{max} values having "tight" control of the spine and possible muscle guarding (van Dieen et al., 2019a, 2019b).

Patients with LBP have demonstrated greater initial LDS (lower λ_{max}) during lifting at the hip in the frontal and transverse planes

* Corresponding author.
E-mail address: rgraham@uottawa.ca (R.B. Graham).

https://doi.org/10.1016/j.jbiomech.2020.110080
0021-9290/

with respect to controls (Asgari et al., 2017) and lower λ_{max} over longer time intervals during repetitive trunk flexion/extension movements (Asgari et al., 2015). Additionally, kinematic variability can be assessed by taking the mean of the standard deviation (MeanSD) of angular displacement at each time point across cycles (Asgari et al., 2015; Graham et al., 2012a, 2012b). Spine kinematics, including angular displacement and velocity, is thought to differ between patients with LBP and controls (Beange et al., 2017; Sedrez et al., 2019). Thus, correct characterization of spine movement in clinical practice is essential in informing diagnosis and providing specific intervention strategies for LBP patients.

Although there is a shift toward assessing spine movement quality and control to improve stratification of LBP diagnosis and care, manual testing and visual inspection are often unreliable (Fritz et al., 2007; Hicks et al., 2003; Stolz et al., 2020). On the other hand, using wearables to perform kinematic evaluations of spinal control has demonstrated acceptable reliability (Graham et al., 2012b; Laird et al., 2016). Furthermore, LDS and MeanSD can be calculated with wearable technology, while visual observations are limited to the assessment of gross dynamics like range of motion (Beange et al., 2019b). Primary and allied healthcare professionals have themselves expressed the need for measurement instruments to be able to quantify specific outcome measures in order to objectively assess movement quality in LBP patients (Beange et al., 2017; van Dijk et al., 2017). Inertial measurement units (IMUs) are being recognized as a portable and cost-worthy alternative to traditional movement quality analyses, and have the potential to be introduced into clinical settings for LBP assessment (Ashouri et al., 2017). The progression towards the widespread use and acceptance of IMU-based assessments in routine clinical practice is currently limited by uncertainties regarding sensor accuracy and reliability (Bauer et al., 2015; Bolink et al., 2016; Cuesta-Vargas et al., 2010; Whelan et al., 2016). Nonetheless λ_{max} and MeanSD have demonstrated moderate to good between-day reliability in healthy participants performing repetitive flexion when calculated with IMU data (Graham et al., 2012b). Moreover, recent work demonstrated good agreement between optical motion capture and IMU data in assessing dynamic spine motor control with LDS and coordination in healthy participants (Beange et al., 2019b).

To support the use of IMUs in research and for them to transition into clinical practice the sensors need to demonstrate good reliability as well as validity, particularly in patient populations. Any evaluation of treatment needs to be considered in the context of measurement error. Reliability is impacted upon by both biological error and error from the equipment used (Hopkins, 2000a). From a biological standpoint, patients with acute LBP have high variability in pain levels day to day (Suri et al., 2011) and pain intensity for those with chronic pain (including LBP) also fluctuates (Peters et al., 2000). Variable levels of pain from one day to another could lead to variable between-day movement patterns. In the case of IMUs there is inherent error from gyroscopic drift and magnetic distortion, and movement amplitude-/frequency-dependent error that needs to be removed with computational models to improve accuracy (Ashouri et al., 2017; Cuesta-Vargas et al., 2010). Examples include improving fusion algorithms to optimally calculate segment and joint orientations and correct for drift (Madgwick et al., 2011; Wittmann et al., 2019). Additionally as the reliability of a given IMU sensor is site and task specific (Cuesta-Vargas et al., 2010), it is necessary to evaluate the reliability of a sensor to be used in an intervention during the specific tasks which will be used to assess the effectiveness of the intervention, as well as in the clinical population of interest, in this case LBP patients.

Therefore, the objective of this study was to assess the between-day reliability of an IMU sensor (HIKOB FOX, Meylan, France) in assessing functional movement quality using LDS and MeanSD in a population with chronic LBP.

2. Methods

2.1. Participants

As part of a larger clinical trial (NCT02059317), thirty patients with LBP (19F, 11M), mean (SD) age of 44 (8) years and BMI of 25.4 (5.6) kg/m^2 were recruited to participate in this study at the University Hospital of Nîmes, France. Participants performed a total of three assessments: the initial baseline assessment, a repeat assessment seven days later, and a third asssement following a five day rehabilition program in a spine rehabilitation centre. The Quebec Back Pain Disability Scale (Kopec et al., 1995) was used for patients to rate their pain from 0 to 100 (with 0 representing no pain and 100 being the worst pain imaginable) on visits one and three. Additionally, patients were asked to rate their pain on a visual analogue scale (0–10) during each of the three tasks on visits one and two (in this case with 0 representing no pain and 10 being the worst pain imaginable). For the purposes of this study, only data from the first two assessments before the intervention were analyzed. All participants provided informed consent prior to data collection and all procedures were approved by the institutional research ethics board (CCP SUD MED III 2013.11.09). Inclusion criteria included: common LBP lasting for more than six months. Exclusion criteria included: previous back surgery, postural disorders, neurological/balance disorders, the use of medications that could impact spine control, and lumbar specific treatment during the last month (infiltration, neurostimulation, patch).

2.2. Instrumentation

HIKOB Fox IMU sensors (Meylan, France) were firmly attached to each participant's back over the T_8 (thorax) and S_2 (pelvis) spinous processes using double-sided tape (Fig. 1D). Raw 3D accelerometer, gyroscope, and magnetometer data were collected at 100 Hz on the data logger embedded in the IMU, which were then downloaded to be analyzed at a later date using custom analysis software.

2.3. Movement protocol

Emulating previous protocols (Bourdon et al., 2019; Dupeyron et al., 2013), on each visit to the lab participants performed 30 repetitions of three different trunk movements in three separate trials (Fig. 1). It has been shown that 30 repetitions is sufficient to achieve acceptable precision of LDS (Dupeyron et al., 2013). The movement tasks were a flexion–extension task (Fig. 1A, representing lifting without a load), a rotation task (Fig. 1B) and a complex task (Fig. 1C, involving movements in three dimensions) and have been described previously (Dupeyron et al., 2013). Briefly, for the flexion–extension task, participants performed trunk flexion from standing to touch a target at knee height with both fingers at one arm length forward. For the rotation task, participants alternated touching targets positioned bilaterally at shoulder height at one arm length laterally on the left and right sides. For the complex task, participants touched targets at knee height on the left, shoulder height on the right, shoulder height on the left and knee height on the right in succession. The mean and SD for the magnitudes of movement for each task in each plane across participants are summarized in the Supplementary Material. The flexion–extension, rotation, and complex tasks were performed to the beat of a metronome at 0.28 Hz, 0.24 Hz, and 0.14 Hz, respectively, based on the preferred pace established in previous work

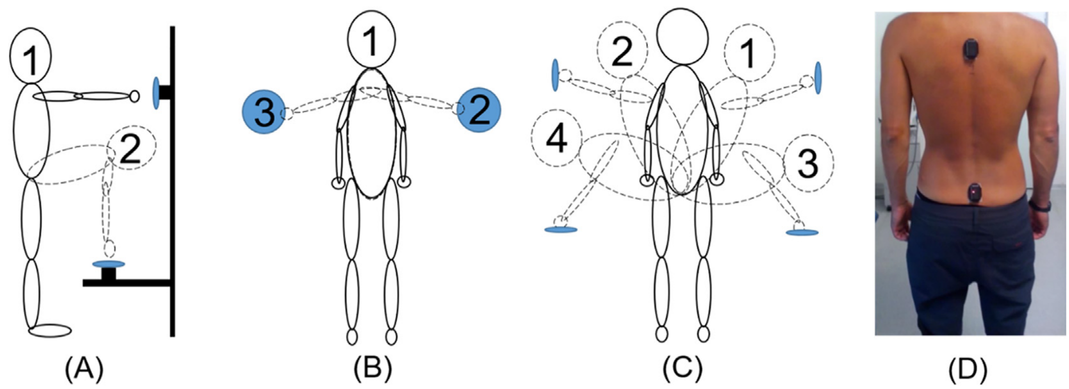

Fig. 1. Outline of each of the movement tasks (adapted with permission from Bourdon et al. 2019). (A) The flexion/extension task in the sagittal plane; (B) the rotation task in the transverse plane and (C) the complex task in three dimensions. The placement of the inertial measurement units (IMUs) is shown in (D).

(Dupeyron et al., 2013). Each movement series was block randomized for each participant, where they each completed the movements in the same order on the two different testing days. The duration of a session was 30 minutes and there were no adverse events during the testing.

2.4. Data processing and analysis

All analyses took place in Matlab R2018b (The MathWorks Inc., Natick, USA). Raw IMU sensor data from each IMU were first fused using the Madgwick algorithm (Madgwick et al., 2011), to compute quaternions and then Euler angles for the pelvis and thorax in global space, as well as for the lumbar spine region (i.e. 3D orientation of the thorax sensor relative to the pelvis sensor). Any drift that was not successfully removed via the sensor fusion process was removed by subtracting a least-squares line of best-fit from the time-series.

2.4.1. Local dynamic stability

To compute LDS for the pelvis, thorax, and lumbar spine using local divergence exponents, two different approaches to state-space reconstruction were taken after selecting a constant number of cycles for each task across participants and then resampling the selected time series to a constant number of samples (300 times the number of cycles) using spline interpolation. First a 6D state space (Y) was created using each independent Euler angle component (x, y, z) at each point in time (t) and one time-lagged version (T_L) as per Equation (1). This was also completed for each independent quaternion component, and while a comparison is presented in the Supplementary Material, only Euler angle results are presented in the manuscript.

$$Y(t) = [x(t) \quad y(t) \quad z(t) \quad x(t+T_L) \quad y(t+T_L) \quad z(t+T_L)] \tag{1}$$

A second state space was created by first calculating the sum of squares (SS) of the three Euler angles (Beange et al., 2019b; Bourdon et al., 2019; Graham et al., 2014; Granata and England, 2006), after first biasing the angles into a positive Cartesian space to remove any zero crossings and relative bias between movement planes (Beaudette et al., 2016). Then a 6D state-space (Y) was reconstructed using the SS (r) at each time point (t) and its time-lagged (T_L) versions as per Equation (2).

$$Y(t) = [r(t) \quad r(t+T_L) \quad r(t+2T_L) \quad r(t+3T_L) \quad r(t+4T_L) \quad r(t+5T_L)] \tag{2}$$

For both state space definitions, a T_L of 10% of the average number of samples per cycle was used (Bourdon et al., 2019; Graham et al., 2014; Granata and England, 2006). The exponential rate of divergence between nearest neighbor trajectories in the reconstructed state space was then determined to estimate the local divergence exponent (λ_{max}) (Bourdon et al., 2019; Dupeyron et al., 2013). This was done by estimating a line of best-fit across the first 0.25 cycles of the average logarithmic divergence curve using both state spaces, using the robust modified Rosenstein algorithm proposed by Mehdizadeh (2019).

2.4.2. Variability

To calculate the variability of the independent Euler angles (x, y, z – in radians) as well as the SS, the MeanSD was used (Bruijn et al., 2009; Graham et al., 2012a). Each angle (i.e. x, y, z, SS) from each set of dynamic trunk movement cycles (i.e. flexion/extension, rotation, complex) was first divided into separate cycles. Each cycle was then time normalized to 101 samples (0–100% of the movement cycle); point by point standard deviations were calculated across all cycles and then the mean value was taken.

2.4.3. Statistics

Reliability was assessed using the intra-class correlation coefficient (ICC $(_{2,1})$), the standard error of measurement (SEM), and the coefficient of variation (CV) in SPSS 26 (IBM Corporation, Armonk, USA). The SEM was calculated as $(SD\Delta/\sqrt{2})$ and CV was calculated as (SEM/Grand Mean) * 100 (Batterham and George, 2003; Hopkins, 2000a). The coefficient of variation is often considered the best measure of reliability, as it allows direct comparison between measures regardless of scaling (Hopkins, 2000a). The minimal detectable difference (MDD) was determined to be 1.5*SEM (Hopkins, 2000a).

3. Results

3.1. Pain scores

The mean (SD) Quebec Back Pain Disability Scale score was 46 (17) on visit one. Pain scores were collected at baseline prior to beginning the movement protocol and during the tasks. The mean (SD) VAS back pain at baseline was 55 (22) and VAS during tasks were available for only 18 observations (out of a possible 90; 30 participants multiplied by three tasks). There was a median decrease of −2, −3 and −1, from visit one to visit two in the VAS for the flexion, rotation and complex tasks, respectively. The change in VAS scores between visits ranged from 0 to −6, suggesting some decrease in pain between visits.

3.2. Local dynamic stability (λ_{max})

There was excellent agreement between λ_{max} calculated using Euler angles and Quaternions ($r \geq 0.99$ for all tasks, Supplementary Material), as such, all subsequent data are presented using the more common Euler angle convention. The mean λ_{max} values for each segment at visit one and visit two in each task are presented in Fig. 2, along with the associated CV. The ICC, SEM and MDD for λ_{max} are presented in Table 1. The average CV for all λ_{max} values was ~10% and all CV values were <20%. Considering CV, reliability was similar between the pelvis and thorax (CV: ~8% and ~10% for xyz and SS respectively) and worse for lumbar motion (CV: ~10% and ~15% for xyz and SS respectively). The CV was generally greater (worse reliability) in the complex task than in the flexion/extension and rotation tasks. ICC values for λ_{max} ranged from 0.28 to 0.81. ICC values across tasks were again generally better for the pelvis and thorax than the lumbar spine (mean 0.64 and 0.73 versus 0.50, respectively). ICC values for the pelvis and thorax were in fact better for the complex task (~0.80) than the flexion/extension task (~0.60) and rotation task (~0.70). However, the ICC of the lumbar spine λ_{max} in the complex task was poor (0.28 [−0.62–0.68]) or moderate (0.53 [−0.06-0.79]), for the xyz and SS respectively. The MDD of λ_{max} ranged from 0.17 to 0.75.

3.3. Variability (MeanSD)

The average MeanSD values for each segment at visit one and visit two in each task are presented in Fig. 3, along with the associated CV. In this case, the MDD ranged from 0.005 rad (0.3°) to 0.032 rad (1.8°). The ICC, SEM and MDD for MeanSD are presented in Table 2. The CV for MeanSD was typically in the range of 20–30%, with the exception of rotations in the frontal plane around the anterior-posterior (AP) axis during the flexion/extension task (33–52%). ICC values for MeanSD ranged from 0.15 to 0.88. ICC values were lowest about the AP axis in the flexion/extension task (ICC: 0.15–0.29) and largest about the longitudinal axis in the rota-

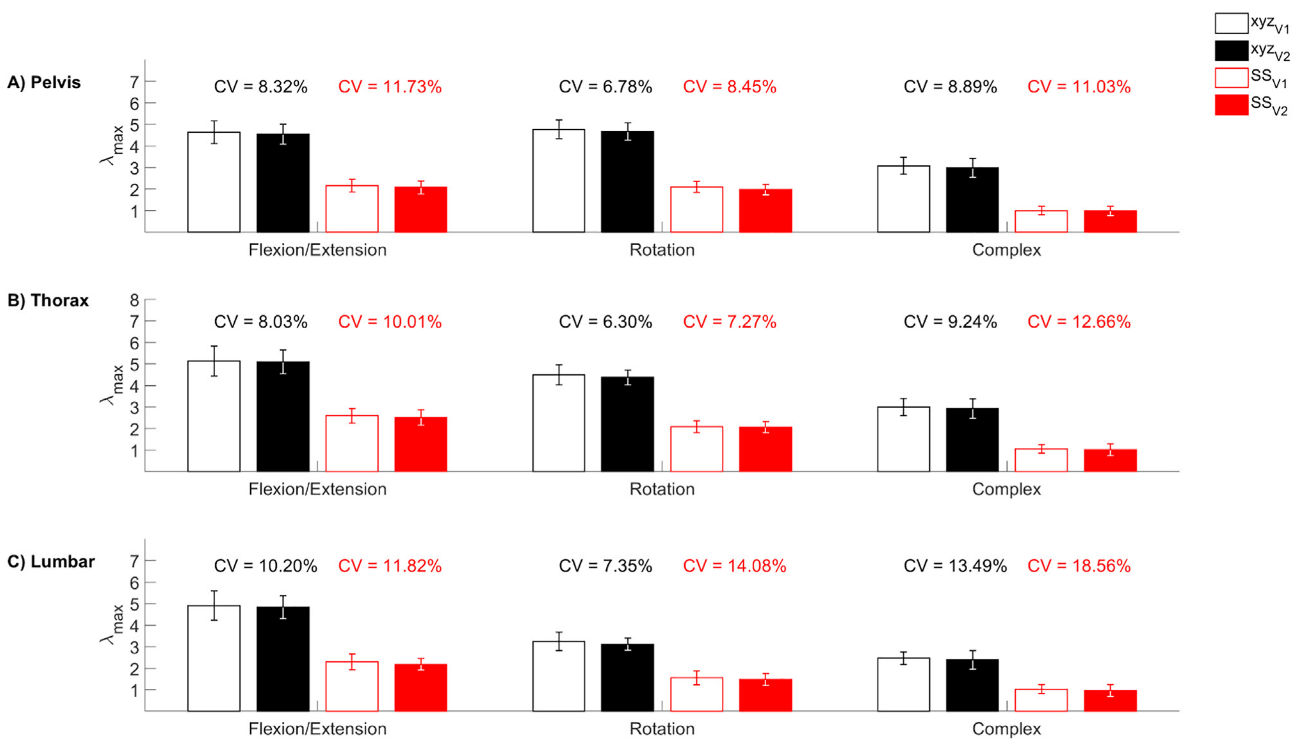

Fig. 2. Local dynamic stability (λ_{max}) on visit one (V1) and visit two (V2) for: A) Pelvis; B) Thorax; and C) Lumbar segments during each task as measured using all xyz angles and the sum of squares (SS). CV = coefficient of variation.

Table 1
Reliability statistics for local dynamic stability (λ_{max}) for each segment during each task.

	Flexion/Extension			Rotation			Complex		
	Pelvis	Thorax	Lumbar	Pelvis	Thorax	Lumbar	Pelvis	Thorax	Lumbar
xyz									
ICC [95% CI]	0.55 [−0.02–0.80]	0.71 [0.34–0.87]	0.49 [−0.16–0.78]	0.62 [0.14–0.83]	0.70 [0.33–0.87]	0.59 [0.08–0.82]	0.75 [0.45–0.89]	0.75 [0.45–0.89]	0.28 [−0.62–0.68]
SEM	0.38	0.41	0.50	0.32	0.28	0.23	0.27	0.27	0.33
MDD	0.57	0.62	0.75	0.48	0.42	0.35	0.40	0.41	0.49
SS									
ICC [95% CI]	0.48 [−0.18–0.77]	0.63 [0.16–0.84]	0.48 [−0.18–0.77]	0.66 [0.25–0.85]	0.78 [0.55–0.91]	0.60 [0.10–0.82]	0.78 [0.51–0.90]	0.81 [0.59–0.92]	0.53 [−0.06–0.79]
SEM	0.25	0.26	0.27	0.17	0.15	0.23	0.11	0.13	0.18
MDD	0.37	0.38	0.40	0.26	0.23	0.32	0.17	0.20	0.28

ICC = intra-class correlation coefficient, SEM = standard error of measurement, MDD = minimal detectable difference, SS = sum of squares, CI = confidence interval.

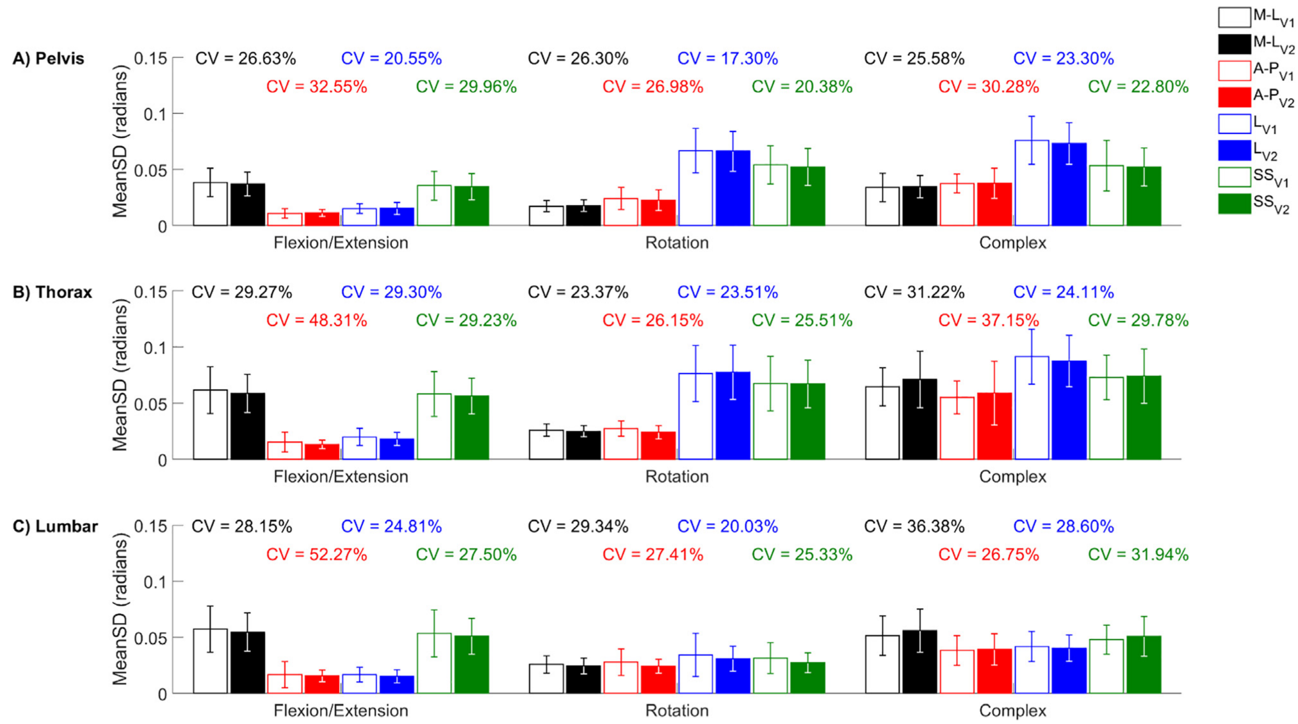

Fig. 3. Variability (MeanSD in radians) on visit one (V1) and visit two (V2) for: A) Pelvis; B) Thorax and C) Lumbar segments during each task as measured about each axis and as a sum of squares. CV = coefficient of variation; M-L = medial–lateral (i.e. flexion/extension); A-P = anterior-posterior (i.e. lateral bending); L = longitudinal (i.e. axial rotation); and SS = sum of squares.

Table 2
Reliability statistics for variability (MeanSD) for each sensor/segment during each task.

	Flexion/Extension			Rotation			Complex		
	Pelvis	Thorax	Lumbar	Pelvis	Thorax	Lumbar	Pelvis	Thorax	Lumbar
Medial-lateral axis (x – Flexion/Extension)									
ICC [95% CI]	0.50	0.38	0.49	0.60	0.50	0.82	0.70	0.39	0.29
	[−0.13–0.78]	[−0.41–0.73]	[−0.16–0.78]	[0.11–0.82]	[−0.12–0.77]	[0.60–0.92]	[0.33–0.87]	[−0.39–0.73]	[−0.62–0.69]
SEM	0.010	0.018	0.016	0.005	0.006	0.007	0.009	0.021	0.019
MDD	0.015	0.026	0.024	0.007	0.009	0.011	0.013	0.032	0.029
Anterior-posterior axis (y – Lateral Bending)									
ICC [95% CI]	0.17	0.15	0.29	0.77	0.47	0.71	0.22	0.32	0.72
	[−0.89–0.63]	[−0.93–0.63]	[−0.62–0.69]	[0.49–0.90]	[−0.19–0.76]	[0.35–0.87]	[−0.78–0.66]	[−0.54–0.70]	[0.37–0.88]
SEM	0.004	0.007	0.009	0.006	0.007	0.007	0.011	0.021	0.010
MDD	0.005	0.010	0.013	0.009	0.010	0.011	0.017	0.032	0.016
Longitudinal axis (z – Axial Rotation)									
ICC [95% CI]	0.70	0.58	0.75	0.86	0.76	0.88	0.64 [0.17–0.84]	0.65 [0.21–0.85]	0.71
	[0.32–0.87]	[0.04–0.81]	[0.43–0.89]	[0.69–0.94]	[0.47–0.89]	[0.74–0.95]			[0.33–0.87]
SEM	0.003	0.006	0.004	0.011	0.018	0.007	0.017	0.022	0.012
MDD	0.005	0.008	0.006	0.017	0.027	0.010	0.026	0.032	0.018
Sum of squares									
ICC [95% CI]	0.47	0.38	0.54	0.81	0.70	0.79	0.83 [0.61–0.92]	0.32	0.40
	[−0.19–0.77]	[−0.42–0.72]	[−0.05–0.80]	[0.57–0.91]	[0.33–0.87]	[0.53–0.91]		[−0.55–0.70]	[−0.36–0.74]
SEM	0.011	0.017	0.014	0.011	0.017	0.007	0.012	0.022	0.016
MDD	0.016	0.025	0.022	0.016	0.026	0.011	0.018	0.033	0.024

ICC = intra-class correlation coefficient, SEM = standard error of measurement, MDD = minimal detectable difference, CI = confidence interval.

tion task (ICC: 0.76–0.88). In the flexion/extension task ICC values about the longitudinal axis (0.58–0.75) were also greater than values about the medial–lateral (ML) axis (0.38–0.50).

4. Discussion

The objective of this study was to assess the between-day reliability of an IMU in assessing spine control and functional movement quality as characterized by λ_{max} and MeanSD in individuals with chronic LBP. Considering the CV, overall reliability was better for λ_{max} than MeanSD (CV: around 10% and 28%, respectively). The ICC values for λ_{max} were generally moderate to good considering Cohen's thresholds for interpreting effect sizes, although no ICC value was above 0.9, which has also been argued as the necessary level to consider a measure practical and valid (Cohen, 1988; Hopkins, 2000b). Unfortunately, we were not able to account for subgroups of LBP within this study, which would give an indication of the heterogeneity of the sample, and may have affected reliability. However, patients with LBP are known to exhibit heterogeneity (van Dieen et al., 2019a) and ICC values increase with increased heterogeneity (Hopkins, 2000a). Additionally, although an apparent decrease in VAS pain scores during movement tasks was recorded between visit one and visit two, the results were based on a very limited number of values (18/90). As such, we cannot realistically account for any variations in pain between visits one and two which may have affected reliability and as such we were not able to use a mixed effect model controlling for pain to estimate reliability. We might expect a difference in pain scores between visits one and two as self-reported pain scores are likely to vary with time (Turk and Marcus, 1994). We also did not account for height or weight in the model, however these anthropometrics have been shown to have only a small correlation with spinal motion and LBP (Heuch et al., 2015; Mellin, 1987), which was also the case with our data here.

The consistency of IMUs with optical motion capture has been shown to be best in the primary direction of trunk movement (Beange et al., 2019a, 2019b, 2018). Consequently, the finding of poorest reliability of MeanSD about the AP axis in the flexion/extension task (ICC: 0.15–0.29, CV: 32.55–52.27%) and the best reliability about the longitudinal axis in the rotation task (ICC: 0.76–0.88, CV: 17.30–23.51%) is not surprising. Greater reliability about the longitudinal axis (ICC: 0.58–0.75, CV: 20.55–29.30%) than the ML axis (0.38–0.50, CV: 26.63–29.27%) in the flexion/extension task was unexpected, although it remains to be seen whether such a difference in the reliability between axes is meaningful. Nevertheless, the reliability of MeanSD calculated using SS was driven by the reliability in the primary axis of movement in the flexion/extension and rotation tasks, which is also similar to earlier work (Beange et al., 2019a, 2019b).

Clinical assessments of spine movement quality and variability are typically performed around the ML axis in the sagittal plane (Delitto et al., 2012; Wattananon et al., 2017). As such moderate reliability during the flexion/extension task for λ_{max} and MeanSD could offer some improvement upon current segmental mobility testing and intervertebral motion testing which have been shown to have poor reliability (Hicks et al., 2003; Stolz et al., 2020). The complexity of processing IMU data was previously a barrier to the uptake of wearables in the clinical setting (Papi et al., 2017). However, new, low cost, wearable sensors synced with mobile devices and cloud-based applications are being developed that provide on-board calculations and signal quality checks in real time (e.g. Graham and Josan, 2017). Additionally, advances in cloud-computing allow data to be stored in secure databases to build large data sets that enable machine learning algorithms to differentiate groups, such as healthy and LBP patients, and provide feedback to the clinician in an easily interpretable format. The MDD presented in this study provide a threshold for both researchers and clinicians to determine whether change has occurred as a result of an intervention. The excellent agreement in λ_{max} when calculated with Euler angles and quaternions, supports the use of the more familiar Euler angle convention in future work for ease of interpretation, although quaternions have the benefit of not suffering from Gimbal lock, should this be present in the dataset.

In conclusion, this study found that LDS (λ_{max}) generally had greater between-day reliability than movement variability (MeanSD) when assessing spine movement using IMUs in patients with LBP.

Acknowledgements

The authors would like to thank the University Hospital of Nîmes for their funding, as well as Dr. Joanna Reeves for her help in preparing the manuscript.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jbiomech.2020.110080.

References

Andersson, G.B.J., 1999. Epidemiological features of chronic low-back pain. Lancet 354, 581–585.
Asgari, M., Sanjari, M.A., Mokhtarinia, H.R., Sedeh, S.M., Khalaf, K., Parnianpour, M., 2015. The effects of movement speed on kinematic variability and dynamic stability of the trunk in healthy individuals and low back pain patients. Clin. Biomech. 30, 682–688.
Asgari, N., Sanjari, M.A., Esteki, A., 2017. Local dynamic stability of the spine and its coordinated lower joints during repetitive Lifting: Effects of fatigue and chronic low back pain. Hum. Mov. Sci. 54, 339–346.
Ashouri, S., Abedi, M., Abdollahi, M., Manshadi, F.D., Parnianpour, M., Khalaf, K., 2017. A novel approach to spinal 3-D kinematic assessment using inertial sensors: towards effective quantitative evaluation of low back pain in clinical settings. Comput. Biol. Med. 89, 144–149.
Balagué, F., Mannion, A.F., Pellisé, F., Cedraschi, C., 2012. Non-specific low back pain. The Lancet 379, 482–491.
Batterham, A.M., George, K.P., 2003. Reliability in evidence-based clinical practice: a primer for allied health professionals. Phys. Ther. Sport 4, 122–128.
Bauer, C.M., Rast, F.M., Ernst, M.J., Kool, J., Oetiker, S., Rissanen, S.M., Suni, J.H., Kankaanpää, M., 2015. Concurrent validity and reliability of a novel wireless inertial measurement system to assess trunk movement. J. Electromyogr. Kinesiol. 25, 782–790.
Beange, K., Chan, A., Graham, R., 2019a. Wearable sensor performance for clinical motion tracking of the lumbar spine. CMBES Proc. 42.
Beange, K.H., Chan, A.D., Beaudette, S.M., Graham, R.B., 2019b. Concurrent validity of a wearable IMU for objective assessments of functional movement quality and control of the lumbar spine. J. Biomech. 97, 109356.
Beange, K.H.E., Beharriell, T.H., Wai, E.K., Graham, R.B., 2017. A spinal control approach to back pain for the primary care provider. J. Curr. Clin. Care 7, 1–14.
Beange, K.H.E., Chan, A.D.C., Graham, R.B. 2018. "Evaluation of wearable IMU performance for orientation estimation and motion tracking," 2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Rome, pp. 1-6, https://doi.org/10.1109/MeMeA.2018.8438623.
Beaudette, S.M., Howarth, S.J., Graham, R.B., Brown, S.H.M., 2016. On the use of a Euclidean norm function for the estimation of local dynamic stability from 3D kinematics using time-delayed Lyapunov analyses. Med. Eng. Phys. 38, 1139–1145.
Bolink, S., Naisas, H., Senden, R., Essers, H., Heyligers, I.C., Meijer, K., Grimm, B., 2016. Validity of an inertial measurement unit to assess pelvic orientation angles during gait, sit-stand transfers and step-up transfers: comparison with an optoelectronic motion capture system. Med. Eng. Phys. 38, 225–231.
Bourdon, E., Graham, R.B., van Dieen, J., 2019. A comparison of methods to quantify control of the spine. J. Biomech. 96, 7.
Bruijn, S.M., van Dieën, J.H., Meijer, O.G., Beek, P.J., 2009. Is slow walking more stable?. J. Biomech. 42, 1506–1512.
Cohen, J., 1988. Statistical power analysis for the behavioral sciences. Lawrence Earlbaum Associates Publisher, Hillsdale, NJ, pp. 8–16.
Cuesta-Vargas, A.I., Galan-Mercant, A., Williams, J.M., 2010. The use of inertial sensors system for human motion analysis. Phys. Ther. Rev. 15, 462–473.
Delitto, A., George, S.Z., Van Dillen, L., Whitman, J.M., Sowa, G., Shekelle, P., Denninger, T.R., Godges, J.J., 2012. Low back pain clinical practice guidelines linked to the international classification of functioning, disability, and health from the orthopaedic section of the American Physical Therapy Association. J. Orthop. Sports Phys. Ther. 42, A1–A57.
Dupeyron, A., Rispens, S.M., Demattei, C., van Dieën, J.H., 2013. Precision of estimates of local stability of repetitive trunk movements. Eur. Spine J. 22, 2678–2685.
Fritz, J.M., Cleland, J.A., Childs, J.D., 2007. Subgrouping patients with low back pain: Evolution of a classification approach to physical therapy. J. Orthop. Sports Phys. Ther. 37, 290–302.
Graham, R., Josan, G. 2017. Development of a novel wearable system for the clinical assessment of movement quality and control in low back pain. In: Proceedings of the XXVI Congress of the International Society of Biomechanics, Brisbane, Australia. 2017
Graham, R.B., Oikawa, L.Y., Ross, G.B., 2014. Comparing the local dynamic stability of trunk movements between varsity athletes with and without non-specific low back pain. J. Biomech. 47, 1459–1464.
Graham, R.B., Sadler, E.M., Stevenson, J.M., 2012a. Local dynamic stability of trunk movements during the repetitive lifting of loads. Hum. Mov. Sci. 31, 592–603.
Graham, R.B., Sheppard, P.S., Almosnino, S., Stevenson, J.M., 2012b. Dynamic spinal stability and kinematic variability across automotive manufacturing work shifts and days. Int. J. Ind. Ergon. 42, 428–434.
Granata, K.P., England, S.A., 2006. Stability of dynamic trunk movement. Spine 31, E271.
Granata, K.P., Gottipati, P., 2008. Fatigue influences the dynamic stability of the torso. Ergonomics 51, 1258–1271.

Heuch, I., Heuch, I., Hagen, K., Zwart, J.A., 2015. A comparison of anthropometric measures for assessing the association between body size and risk of chronic low back pain: the HUNT study. PLoS ONE 10, 15.
Hicks, G.E., Fritz, J.M., Delitto, A., Mishock, J., 2003. Interrater reliability of clinical examination measures for identification of lumbar segmental instability. Arch. Phys. Med. Rehabil. 84, 1858–1864.
Hopkins, W.G., 2000a. Measures of reliability in sports medicine and science. Sports Med. 30, 1–15.
Hopkins, W.G., 2000b. A scale of magnitudes for effect statistics. In: A new view of statistics.
Kopec, J.A., Esdaile, J.M., Abrahamowicz, M., Abenhaim, L., Wooddauphinee, S., Lamping, D.L., Williams, J.I., 1995. The Quebec back pain disability scale - measurement properties. Spine 20, 341–352.
Laird, R.A., Kent, P., Keating, J.L., 2016. How consistent are lordosis, range of movement and lumbo-pelvic rhythm in people with and without back pain?. BMC Musculoskeletal Disord 17, 403.
Madgwick, S.O.H., Harrison, A.J.L., Vaidyanathan, R., Ieee, 2011. Estimation of IMU and MARG orientation using a gradient descent algorithm. In: 2011 Ieee International Conference on Rehabilitation Robotics. Ieee, New York.
Mehdizadeh, S., 2019. A robust method to estimate the largest Lyapunov exponent of noisy signals: a revision to the Rosenstein's algorithm. J. Biomech. 85, 84–91.
Mellin, G., 1987. Correlations of spinal mobility with degree of chronic low back pain after correction for age and anthropometric factors. Spine 12, 464–468.
Mokhtarinia, H.R., Sanjari, M.A., Chehrehrazi, M., Kahrizi, S., Parnianpour, M., 2016. Trunk coordination in healthy and chronic nonspecific low back pain subjects during repetitive flexion-extension tasks: effects of movement asymmetry, velocity and load. Hum. Mov. Sci. 45, 182–192.
Murray, C.J., Vos, T., Lozano, R., Naghavi, M., Flaxman, A.D., Michaud, C., Ezzati, M., Shibuya, K., Salomon, J.A., Abdalla, S., 2012. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. The Lancet 380, 2197–2223.
Papi, E., Koh, W.S., McGregor, A.H., 2017. Wearable technology for spine movement assessment: a systematic review. J. Biomech. 64, 186–197.
Peters, M.L., Sorbi, M.J., Kruise, D.A., Kerssens, J.J., Verhaak, P.F.M., Bensing, J.M., 2000. Electronic diary assessment of pain, disability and psychological adaptation in patients differing in duration of pain. PAIN® 84, 181–192.
Rosenstein, M.T., Collins, J.J., De Luca, C.J., 1993. A practical method for calculating largest Lyapunov exponents from small data sets. Physica D 65, 117–134.
Ross, G.B., Sheahan, P.J., Mahoney, B., Gurd, B.J., Hodges, P.W., Graham, R.B., 2017. Pain catastrophizing moderates changes in spinal control in response to noxiously induced low back pain. J. Biomech. 58, 64–70.
Seay, J.F., Van Emmerik, R.E.A., Hamill, J., 2011. Low back pain status affects pelvis-trunk coordination and variability during walking and running. Clin. Biomech. 26, 572–578.
Sedrez, J.A., de Mesquita, P.V., Gelain, G.M., Candotti, C.T., 2019. Kinematic characteristics of sit-to-stand movements in patients with low back pain: a systematic review. J. Manip. Physiol. Ther. 42, 532–540.
Spinelli, B.A., Wattananon, P., Silfies, S., Talaty, M., Ebaugh, D., 2015. Using kinematics and a dynamical systems approach to enhance understanding of clinically observed aberrant movement patterns. Man. Ther. 20, 221–226.
Stolz, M., von Piekartz, H., Hall, T., Schindler, A., Ballenberger, N., 2020. Evidence and recommendations for the use of segmental motion testing for patients with LBP - A systematic review. Musculoskelet. Sci. Pract. 45, 12.
Suri, P., Rainville, J., Fitzmaurice, G.M., Katz, J.N., Jamison, R.N., Martha, J., Hartigan, C., Limke, J., Jouve, C., Hunter, D.J., 2011. Acute low back pain is marked by variability: an internet-based pilot study. BMC Musculoskeletal Disord. 12, 220.
Turk, D.C., Marcus, D.A., 1994. Assessment of chronic pain patients. Semin. Neurol. 14, 206–212.
van Dieen, J.H., Reeves, N.P., Kawchuk, G., van Dillen, L.R., Hodges, P.W., 2019a. Motor control changes in low back pain: divergence in presentations and mechanisms. J. Orthop. Sports Phys. Ther. 49, 370–379.
Van Dieen, J.H., Reeves, P., Kawchuk, G., Van Dillen, L.R., Hodges, P.W., 2019b. Analysis of motor control in patients with low back pain: a key to personalized care?. J. Orthop. Sports Phys. Ther. 49, 380–388.
van Dijk, M., Smorenburg, N., Visser, B., Heerkens, Y.F., Nijhuis-van der Sanden, M. W.G., 2017. How clinicians analyze movement quality in patients with non-specific low back pain: a cross-sectional survey study with Dutch allied health care professionals. BMC Musculoskeletal Disord. 18, 11.
Waddell, G., 2005. Subgroups within" nonspecific" low back pain. J. Rheumatol. 32, 395–396.
Wattananon, P., Ebaugh, D., Biely, S.A., Smith, S.S., Hicks, G.E., Silfies, S.P., 2017. Kinematic characterization of clinically observed aberrant movement patterns in patients with non-specific low back pain: a cross-sectional study. BMC Musculoskeletal Disord. 18.
Whelan, D., O'Reilly, M., Huang, B.Q., Giggins, O., Kechadi, T., Caulfield, B., 2016. Leveraging IMU data for accurate exercise performance classification and musculoskeletal injury risk screening. In: Patton, J., Barbieri, R., Ji, J., Jabbari, E., Dokos, S., Mukkamala, R., Guiraud, D., Jovanov, E., Dhaher, Y., Panescu, D., Vangils, M., Wheeler, B., Dhawan, A.P. (Eds.), 2016 38th Annual International Conference of the Ieee Engineering in Medicine and Biology Society. Ieee, New York, pp. 659–662.
Wittmann, F., Lambercy, O., Gassert, R., 2019. Magnetometer-based drift correction during rest in IMU arm motion tracking. Sensors 19, 1312.
Wu, A., March, L., Zheng, X., Huang, J., Wang, X., Zhao, J., Blyth, F.M., Smith, E., Buchbinder, R., Hoy, D., 2020. Global low back pain prevalence and years lived with disability from 1990 to 2017: estimates from the Global Burden of Disease Study 2017. Ann. Transl. Med. 8, 299.